(12) UK Patent Application (19) GB (11) 2 368 679 (13) A

(43) Date of A Publication 08.05.2002

(21) Application No 0110495.9

(22) Date of Filing 30.04.2001

(30) Priority Data
(31) 09585882 (32) 31.05.2000 (33) US

(71) Applicant(s)
International Business Machines Corporation
(Incorporated in USA - New York)
Armonk, New York 10504, United States of America

(72) Inventor(s)
Gregory M Ball

(74) Agent and/or Address for Service
IBM United Kingdom Limited
Hursley Park, WINCHESTER, Hampshire, SO21 2JN,
United Kingdom

(51) INT CL^7
G06F 17/30

(52) UK CL (Edition T)
G4A AUDB

(56) Documents Cited
US 5787412 A   US 5721911 A

(58) Field of Search
UK CL (Edition T) G4A AUDB
INT CL^7 G06F 17/30
Online: WPI EPODOC JAPIO

(54) Abstract Title
Meta data catalogue

(57) A data warehouse, comprising means for database modelling, and means for extracting data from the database and storing the extracted data. A reporting tool is used to make inquiries of the stored data and to format the results of the inquiries, and a single meta data catalogue is coupled to the means for modelling, the means for extracting and storing, and the reporting tool. With the preferred embodiment of this invention, a user needs to update information only once. The meta data then drives changes in associated systems, such as the warehouse, extraction programs and reporting tools. In order to accomplish this, the meta data catalogue holds enough information to control all these various parts of the process.
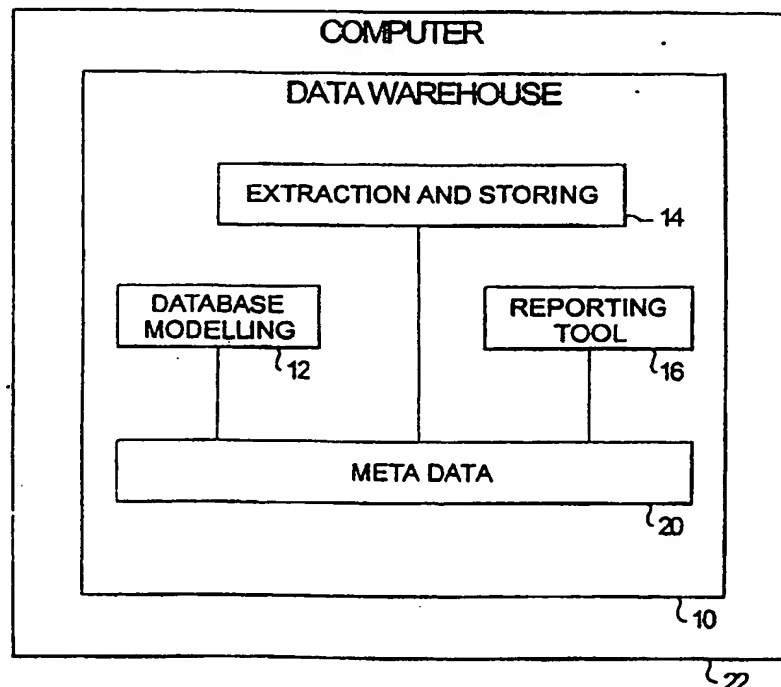
GB 2 368 679 A

1 / 1
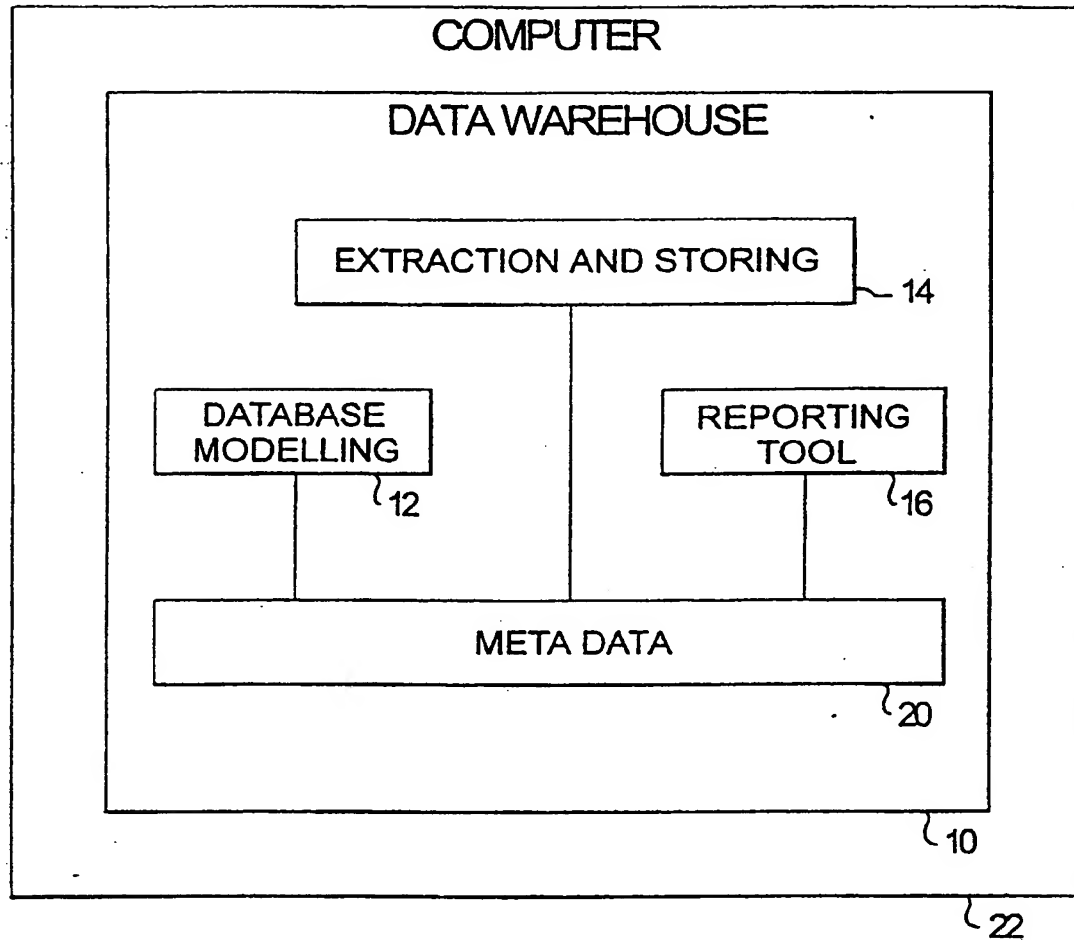
2368679

1

META DATA CATALOGUE

Technical Field of the Invention

5   This invention generally relates to meta data, and more specifically,
to a meta data repository that stores summary information about data.

Background of the Invention

10   Meta data is information about data. For example, meta data about a
particular database might contain descriptions about what type of
information is held in each table and field within the database. The
database's meta data would also explain the physical attributes of the
database such as the types and lengths of the fields, table spacing,
15   indexing, etc. Basically, the meta data repository should be the one place
to find out everything and anything about a database.

          Meta data catalogues are collections or repositories of meta data.
These catalogues are playing an increasingly vital role in publishing
authenticated information and in storing and disseminating such information
20   through a controlled but uniform interface. While these catalogues are of
substantial use, the normal meta data catalogue has several constraints.
For instance, these catalogues are inactive -- that is, nothing can really
be done to the data in the catalogues. Also, these catalogues provide only
limited information. The meta data catalogue may provide complete
25   information about a particular database, but all the information in that
database came from another database, and often a user needs to know more
about that other database.

30   Various attempts have been made to eliminate or alleviate these
constraints. A persistent problem, however, is that the catalogues do not
hold all the information a user might want. In particular, while a meta
data catalogue may hold all the physical information of a database, a user
is normally not able to determine from the catalogue, where the information
35   came from -- that is, the mappings indicating how one database maps to
another.

          In addition, when several systems are used that all need meta data to
function and that all have their own meta data repositories, these
40   repositories must be kept in synchronisation with each other. For
instance, typically in a data warehouse, extraction programs are used to
move data from the operational system to the warehouse, and reporting tools
are used to allow business analysts to analyse the warehoused data. Both
the extraction programs and the reporting tools have their own meta data;

2

and if a field on one table in the operation system changes, this information must be updated in several places: the warehouse itself, the extraction programs, and the reporting tools.

5   It would be desirable to provide a meta data catalogue with the ability to drive changes in associated systems.

DISCLOSURE OF THE INVENTION

10   Accordingly, the present invention provides a data warehouse, comprising means for database modelling, and means for extracting data from the database and storing the extracted data. A reporting tool is used to make inquiries of the stored data and to format the results of the inquiries, and a single meta data catalogue is coupled to the means for modelling, the means for extracting and storing, and the reporting tool.
15 With the preferred embodiment of this invention, a user needs to update information only once. The meta data then drives changes in associated systems, such as a warehouse, extraction programs and reporting tools. In order to accomplish this, the meta data catalogue holds enough information
20 to control all these various parts of the process.

The present invention allows a developer to quickly model a new database based on an existing database, identify how the two databases are linked to each other (mapping information), create the database definition,
25 and then use all this captured information in a variety of reporting methods. All the pertinent data are in one location. The meta data catalogue is of particular use in very large databases .(1000 + tables), because much of the functionality included allows changes to the data model to be made quickly and globally when needed.

30
BRIEF DESCRIPTION OF THE DRAWING

The present invention will now be described, by way of example only, with reference to the following drawing which illustrates a data warehouse
35 embodying the present invention.

DETAILED DESCRIPTION OF THE INVENTION

The Figure shows a data warehouse 10, comprising means 12 for
40 database modelling, and means 14 for extracting data from said database and storing the extracted data. A reporting tool 16 is used to make inquiries of the stored data and to format the results of the inquiries, and a single meta data catalogue 20 is coupled to the means for modelling, the means for extracting and storing, and the reporting tool. Preferably, a user needs

3

to update information only once. The meta data 20 then drives changes in associated systems, such as warehouse 10, extraction programs 14, and reporting tools 16. In order to accomplish this, the meta data catalogue 20 holds enough information to control all these various parts of the process.

5

The meta data repository stores summary information about data. For example, meta data about a particular database might contain descriptions about what type of information is held in each table and field within the database. The database's meta data would also explain the physical attributes of the database. The invention allows the developer to quickly model a new database based on an existing database, identify how the two databases are linked to each other (mapping information), create the database definition, and then use all this captured information in a variety of reporting methods.

10

15

The invention stores meta data from different databases in a single meta data catalogue 20. All the pertinent data is in one location. The meta data catalogue 20 is of particular use in very large databases (1000 + tables), because much of the functionality included allows changes to the data model to be made quickly and globally when needed.

20

The data warehouse 10 may be part of a database computer 22. More specifically, database computer 22 includes one or more storage media containing the data warehouse 10. Database computer 22 is preferably a massively parallel process computer and executes the UNIX operating system or the Windows NT operating system, although other computer and operating system configurations are also envisioned by the present invention (UNIX is a trademark exclusively licensed by the Open Group and Windows NT is a trademark of Microsoft Corporation). Data warehouse 10 is suited to run on any computer that supports an Open Database Connect (ODBC) interface to data warehouse.

25

30

4

CLAIMS

1.   A data warehouse (10) comprising:
        means (12) for database modelling;
5   means (14) for extracting data from said database and storing said
extracted data;
        a reporting tool (16) for making inquiries of said stored data and
formatting the results of said inquiries; and
        a single meta data catalogue (20) coupled to said means for
10   modelling, said means for extracting and storing, and said reporting tool.

2.   A data warehouse (10) according to claim 1, wherein the meta data
catalogue (20) includes means to update information in the database and the
reporting tool (16).

15

3.   A data warehouse (10) according to claim 1, wherein the means (14)
for extracting includes an extraction program, and the meta data catalogue
(20) includes means to update information in the extraction program.

20   4.   A data warehouse (10) according to claim 1, wherein a user needs to
update information only once in the meta data catalogue (20), and the meta
data then drives changes in the means for extracting and the reporting
tool.

25

The Patent Office

INVESTOR IN PEOPLE

Application No:	GB 0110495.9	Examiner:	Nicholas Mole
Claims searched:	1-4	Date of search:	4 March 2002

Patents Act 1977
Search Report under Section 17

Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

    UK Cl (Ed.T): G4A UDB

    Int Cl (Ed.7): G06F 17/30

Other: Online: WPI EPODOC JAPIO

Documents considered to be relevant:

Category	Identity of document and relevant passage	Relevant to claims
A	US 5787412 (BOSCH)	
A	US 5721911 (KHANH)	

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.

An Executive Agency of the Department of Trade and Industry